Integrated multi-domain risk assessment using automated hypothesis testing

Gehrke, Oliver; Heussen, Kai; Korman, Matus

Published in:
Proceedings of the 2nd Workshop on Cyber-Physical Security and Resilience in Smart Grids

Link to article, DOI:
10.1145/3055386.3055398

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Integrated Multi-Domain Risk Assessment Using Automated Hypothesis Testing

Oliver Gehrke
Technical University of Denmark
Center for Electric Power and Energy
4000 Roskilde, Denmark
olge@elektro.dtu.dk

Kai Heussen
Technical University of Denmark
Center for Electric Power and Energy
4000 Roskilde, Denmark
kh@elektro.dtu.dk

Matus Korman
KTH Royal Institute of Technology
Electric Power and Energy Systems
100 44 Stockholm, Sweden
korman@kth.se

ABSTRACT

In this paper we present an approach for the integration of cybersecurity tools from multiple domains into an overall risk assessment framework which takes the complex interactions between domains in smart grid systems into account. The approach is based on generating hypotheses from a template, which are then analyzed for their probability and associated impact on the system. The feasibility of the proposed approach is discussed using a very simple example case to serve as a proof of concept. Furthermore, we introduce a generic software framework for the processing of hypothesis templates.

CCS CONCEPTS

•Hardware →Smart grid; •Security and privacy →Intrusion detection systems; •Theory of computation →*Automated reasoning;*

KEYWORDS

risk assessment, hypothesis testing, cybersecurity, power system, distribution grid, distributed energy resources, intrusion detection

ACM Reference format:
Oliver Gehrke, Kai Heussen, and Matus Korman. 2017. Integrated Multi-Domain Risk Assessment Using Automated Hypothesis Testing. In *Proceedings of The 2nd Workshop on Cyber-Physical Security and Resilience in Smart Grids, Pittsburgh, PA USA, April 2017 (CPSR-SG 2017),* 6 pages.
DOI: http://dx.doi.org/10.1145/3055386.3055398

1 INTRODUCTION

The operation of electrical distribution grids, as seen from the utility control center, is centered around the SCADA and DMS systems in place. In practice this means that the operator in the control center achieves situational awareness about the state of the physical infrastructure through a combination of online monitoring and partly offline decision support tools available in the control center. The severity of threats to power system security of supply is evaluated continuously, quantified by assessment tools that integrate operational data delivered by the SCADA system with background knowledge on the present system configuration. Examples of such tools are state estimators, contingency screening and real-time security indicators.

In a different context, but similar situation, utilities have increasingly been facing cyber-threats and attacks on the ICT backbones of the systems named above. Utilities have increasingly engaged in building cyber-oriented active and passive defense mechanisms and infrastructure, to keep (at least) the central OT safe from external threats. A deeper understanding of possible penetration depth in the utility IT and OT systems can be achieved by architecture assessment systems (such as CySeMoL). Such and other tools facilitate proactive defense-in-depth strategies on the field of cybersecurity threats and countermeasures. However, it is not trivial to assess the potential impact of cyber attacks on the security of supply.

Considering increasing frequency and sophistication of attacks, it cannot be assumed that the central operational systems will always be safe, even with continuously improving cyber defenses. It is evident that cyber attacks aimed at physical supply disruptions are not only imaginable, but a real possibility [23].

The point of departure for the present work is the understanding that utilities in the near future need to be able to assess the potential impact on the security of supply caused by cyber attacks. Furthermore, operators need to respond quickly — ideally in real time — to new threats as they appear. To be able to do this, new tools are needed to map out and prioritize their response to these threats.

This need is becoming more pronounced due to the proliferation of smart grid technologies, i.e. a trend towards a higher degree of automation in the power grid and the deployment of automation solutions in areas of the grid which were previously not connected to communication networks, such as customer meters, building automation systems and the control of distributed energy resources (DER). While power grids and the ICT systems needed for their operation have traditionally been treated as separate infrastructures, smart grids have to be analyzed as cyber-physical systems. Consequently, IT security is only one aspect of the cybersecurity of smart grids.

The SALVAGE project aims at developing and integrating analysis methods which are behind three domain-specific tools. They provide a quantification of cybersecurity vulnerability, probability of an ongoing intrusion and power system impact. One key challenge is how to combine these three types of information, which are of very diverse nature, into an integrated analysis of the present system state. Another challenge is how to go about stepwise automation of such an analysis, in the light of the very large amounts of data which need to be analysed.

CPSR-SG 2017, Pittsburgh, PA USA

DOI: http://dx.doi.org/10.1145/3055386.3055398

In this paper, we present an approach based on automated hypothesis testing. We discuss the premises and core concepts of the chosen approach, then present an example for a very simple cyber-physical system as a proof of concept. Finally, we discuss the implementation of the concept in a software framework, the development of which has been one of the activities in the SALVAGE project. We conclude with a brief discussion of the next steps to be taken.

2 RELATED WORK

The cyber-physical systems perspective has received increasing attention from the smart grids community [10, 17]. The range of approaches includes purely analytical methods to attack detection [13], system modelling and (co-)simulation [9], formal methods for security assessment [3] or test specification [21].

Anomaly detection identifies rare data instances or events that do not match an expected pattern [2]. The development of models used for anomaly detection requires domain expertise, and additionally, if data driven models are required, data analysis knowledge. Model-based approaches for anomaly detection can be based on Bayesian networks [7]. Once both cyber and physical anomaly detection analysis is performed, cyber-physical metrics need to be developed to combine the information from both domains and address the tight relations between the power system and the ICT domains. Anomaly detection with regression models has been used for discovering cyber attacks on a SCADA system [20], wind turbine fault detection [22] and photovoltaic plant fault diagnostics [14]. A special case of an attack against voltage control in distribution power grids has been described in [5].

A number of methods can be used to assess the cyber-vulnerability of IT and OT architectures such as those enabling the automated control and operation of the smart grid. Examples include methods for traditional human-performed information security risk asseessments (e.g., FAIR [18] or HMG IA [11]), and risk assessments tailored for the smart grid (e.g., SGIS Toolbox [1]). There also exist approaches and tools to automate vulnerability assessment, mostly based on attack graphs [6, 15], such as MulVAL [12], ADVISE [8], CyberSAGE [19] or CySeMoL [4, 16]. Many of the established and standardized methods for risk assessment are being used in the industry, and can lead to accurate results when used by skilled professionals. However, their use requires much effort each time an assessment is being performed. The attack graph based approaches are less labor-intensive, since much of the analysis is performed by a computer; however, tend to require a human to formulate the ontology of how attacks may happen (i.e., attack steps, transitions between them, and dependencies of the transitions). CySeMoL has the advantage of being such an ontology, aiming to be comprehensive, generic, and even having undergone scientific scrutiny. Hence, it has been chosen as one of the tools used by the framework presented below.

3 PROPOSED APPROACH

Traditionally, and from the point of view of a control room operator in an electrical power system, the ICT aspect of the power system (which includes the SCADA and DMS systems) has not been part of the system to be operated; it was an infrastructure which was assumed to work. Today, the ICT domain (both OT and IT) is explicitly taken into account during operations. However, the pure ICT risk assessment still tends to separate security relevant events from ICT and physical domains. Our approach assumes an operational context in which the risk of several cyber-security breaches are evaluated at the same time, and where there is highly uncertain information about possible security breaches. In such a context only an integrated assessment is meaningful, where a risk-oriented prioritization of potential threats and impacts is required to accommodate probabilistic information.

3.1 Background

Cross-dependencies between domains make it hard to isolate the impact analysis of the physical and cyber domains. Consequences of IT-domain breaches which manifest themselves in the physical domain, are not quantifiable using the same metrics as a pure analysis of the IT domain. Furthermore, the model types and propagation mechanisms are different in each of these domains. Different propagation mechanisms make a direct model integration impossible; a coupling of models could however be achieved, similar to the strategy for co-simulation approaches. A pure co-simulation based assessment would however require a full model parametrization, simulation of combined models and result assessment. Probabilistic input hypotheses and dependencies would still not be feasible in such an approach. Nevertheless, the good success of probabilistic modeling in – for example – intrusion detection or ICT architecture assessment (CySeMol), indicates that uncertain knowledge from several domains can successfully be combined. Attack modeling methods based on probabilistic methods and tree structure are quite successful [6].

However, attack probability alone does not constitute sufficient information if the goal is to be able to prioritize the response to several attack scenarios. The impact quantity, i.e. the consequence of an attack in terms of the physical system, is of equal importance. Calculating the overall risk of a particular attack scenario (as a function of probability and potential impact) cannot be done without integrating an impact analysis into the vulnerability analysis process, which is necessarily domain-specific.

3.2 Problem Statement: Key requirements

We see a need for a framework that performs a real time (online) integrated assessment of the state of the power grid, using aspects from multiple domains (power system impact, intrusion detection, cyber vulnerability) and multiple sources of input (power system measurements, distributed energy resources (DERs), IT and OT systems) – a framework that aims to provide a joint prioritization of possible threat/impact-scenarios, taking into account uncertainty of input information.

We see that an automated, integrated assessment has the potential to yield more accurate results than performing several separate assessments whose results need to be interpreted and finally merged and synthesized by a human. The strength of the integrated assessment resides in (1) the possibility to handle dependencies that cross boundaries of different domains in an automated fashion and already during the assessment process; and hence (2) the ability of the framework to operate in real time (online assessment), and

therefore the ability to provide frequent re-assessments based on an actual state of the power grid and its IT/OT infrastructure.

3.3 Core Concepts

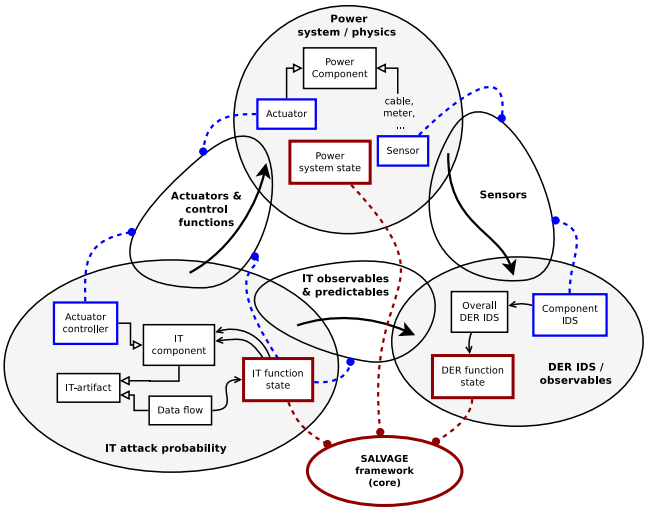

Figure 1: Domain interactions

The *System configuration* is the model of the underlying system that is being evaluated – including the power grid, distributed energy resources attached to it, and all relevant pieces of IT/OT infrastructure that govern and/or otherwise interact with any and all of the former.

The term *Domain*, in the context of this paper, is referring to one of the following: cyber security/vulnerability, DER intrusion detection, or power system [impact] analysis. Figure 1 shows the relationships between the domains, expressed in terms of the possible flows of information between them.

A *Domain-specific assessment* (DSA) is a quantification (function) obtained from a computational component. For a given input parametrization, the DSA quantifies *one* or *several* (branching!) pairs of a *probability* (p) and a *different quantity* (a Domain specific variable, e.g. electric current, operating state, switch state). The input parametrization may include both static parameters (e.g. system configurations), dynamic online data (e.g. meter readings), and other [dynamic] invocation parameters (e.g. state of a particular switch or function). Example: e.g. given a certain state to occur (DER on or off), what is the quantitative impact (power distribution), and what is the chance of it occcuring.

Domain specific variable (DSV) is a quantity (an intermediate or final impact value, e.g. electric current) that is specific to a DSA, and can be obtained by the DSA, and may be a (partial) input parametrization of another DSA.

A *Hypothesis template* is a tree-like structural representation (strictly taken, an acyclic directed graph) of the following:

(1) An "attack tree" in form of a probability tree using logical gates, which aggregates probabilistic data from DSA inputs at its leaf level all the way up to the top-level probability of the assessed realism of a hypothesized scenario;
(2) A set of domain-specific variables (DSVs, e.g. power supply loss due to opening of a specific breaker) related to a system configuration – the "output quantities" of DSAs, which, according to the structure of the hypothesis template, ultimately contribute to the risk value (the aggregated product of partial probabilities and impacts) of the hypothesized scenario. The DSVs can be defined together with an "input parameterization" of a DSA, since the DSA might need to be dynamically invoked with the output parameters of another DSA (instead of the baseline system configuration);
(3) The coupling of output probabilities of each DSA to the probability tree (through logical gates such as AND, OR, XOR and NOT);
(4) The coupling of DSA outputs with invocations of other DSAs (at a higher level in the tree), leading to a final scenario quantification.

The *Risk node* is a node used as the root node of a hypothesis template, joining the final impact quantity (associated with an attack goal) with the final hypothesis probability into a single, overall risk value of the (grounded) hypothesis.

A *Risk belief* is the quantified risk value of a hypothesis (hypothesized scenario).

The *Hypothesis probability* (belief) is a fully quantified probability value of a hypothesis.

Figure 2 illustrates the concepts defined above and their relations. In a hypothesis, probabilities and other quantities are aggregated through functional gates all the way up to the root value of a hypothesis which represents its risk value, i.e. the product of the impact of an adverse event and its probability.

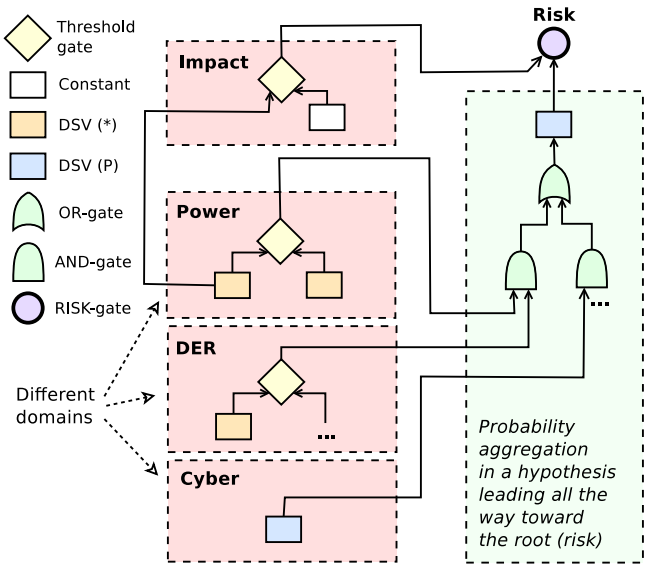

Figure 2: Conceptual structure of a hypothesis template and resolution

4 PROOF OF CONCEPT

4.1 The fuse blowing scenario

In the following, we will demonstrate the feasibility of the above approach. As discussed in section 3.1, (say something about that the thing to be demonstrated is not the ability of the method to handle large power systems but multiple domains). We use the test system shown in figure 3 which represents a multi-domain system while keeping the complexity within the individual domains at an absolute minimum, i.e. a simple grid configuration, simple IT network and simple DER behaviour. The test grid configuration consists of a PV inverter connected to an electrical power grid of infinite capacity. The inverter can be remote controlled from a laptop (equipped with remote-control software) through a minimalistic communication network consisting of a direct, unswitched ethernet connection between a laptop and the controllable inverter. A single inverter function is available through the remote-control interface: Switching the inverter on — which would cause the inverter to feed power to the grid, depending on current solar irradiation — or off, which would reduce the current flowing between inverter and grid to exactly zero. The cable between inverter and grid is protected by an infinitely small fuse (here: 1μA) , such that the fuse would inevitably blow and disconnect the inverter from the grid as soon as the inverter is switched on. This reduces the range of possible grid impacts of a cyberattack to a binary choice: Turning the inverter on will result in permanent damage to the system; the associated need for repair is easily quantifiable in terms of financial damage. The only alternative action is to leave the inverter in the off state, which has no consequences.

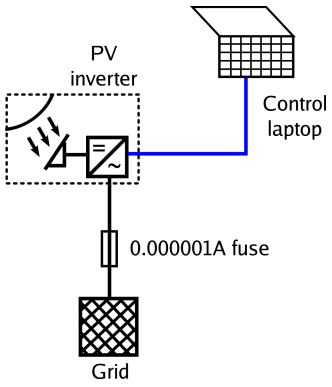

Figure 3: Fuse blow scenario

4.2 Processing the fuse blow scenario

Applying the method described in section 3.3 results in the hypothesis template shown in figure 4.

Starting at the bottom of the tree, the first two inputs to the resolver are the outputs of the DER intrusion detection tool [DER-IDS] and the vulnerability analysis tool [IT]. Combined, they express the probability of the inverter having been compromised. At the next level, the actions of the attacker must be considered; specifically what the attacker intends to do with the compromised system. Due to the design of the test case, the attacker has only two options available: switch the inverter on, or leave it off. These two options, while the probabilities for their respective occurrences may be known, lead to a widely different outcome for the physical system. Therefore, two different hypotheses must be generated from the hypothesis template, and each outcome must be evaluated separately. For each hypothesis, the response of the inverter to attacker behaviour must be evaluated using a component model of the inverter [DER-CM]. The output of the component model is twofold and consists of (a) the *magnitude* of the electrical current at the inverter terminals in case of the inverter switching on, and (b) the *probability* of the inverter switching on after receiving a remote switching command. The current magnitude is then inserted into a system model [PS] which calculates the line flow between inverter and grid connection, thereby determining the current through the fuse. By comparing this result with the fuse characteristics, it can be transformed into the probability of an inverter operation causing a fuse blow, which in turn can be combined with the probability of an inverter operation to yield the probability of a fuse blow. Independent of the latter, deduction-based result, a blown fuse may also be observed by a system monitoring the state of the physical system in real time [SSM]. The disjunction of both yields the overall probability of a blown fuse. Multiplication with the calculated impact — here expressed as the financial damage caused by the required replacement of the fuse — yields the risk value.

The scenario presented above is very simplistic. Nevertheless, the application of more complex scenarios is not expected to lead to notably greater tree depths in the hypothesis template. Rather, they will lead to increased width of the trees, as the multiplicity of components in a physical system is taken into account. A practical assessment system would be built around a library of hundreds or thousands of different hypothesis templates and perform re-evaluation on a continuous basis. This would provide a human operator or another information system with a frequently updated set of the most threatening scenarios according to the present state of the power grid and its cyber infrastructure.

5 PRACTICAL IMPLEMENTATION

As part of the SALVAGE project, the process of hypothesis generation (expansion) from a hypothesis template, as well as the evaluation and ranking of multiple hypotheses, has been automated by developing a generic framework. Figure 5 shows the high-level architecture of this framework. A hypothesis template (represented by a hierarchically structured file) is given as input to a hypothesis generator. The generator analyzes the template and extracts a list of branch points which require expansion into multiple hypotheses. Furthermore, the generator identifies the type of information represented by the root node and instantiates an appropriate ranking algorithm for this type of node. As a next step, the generator expands the template at all n_b branch points by filling in one of the possible states at each branch point, thus sequentially generating 2^{n_b} hypotheses. Currently, this simple brute-force approach to hypothesis generation/expansion is the only available strategy. Each hypothesis is being forwarded to a resolver module which must match the hypothesis template. The resolver module traverses the tree and identifies the dependencies between ordinary tree nodes

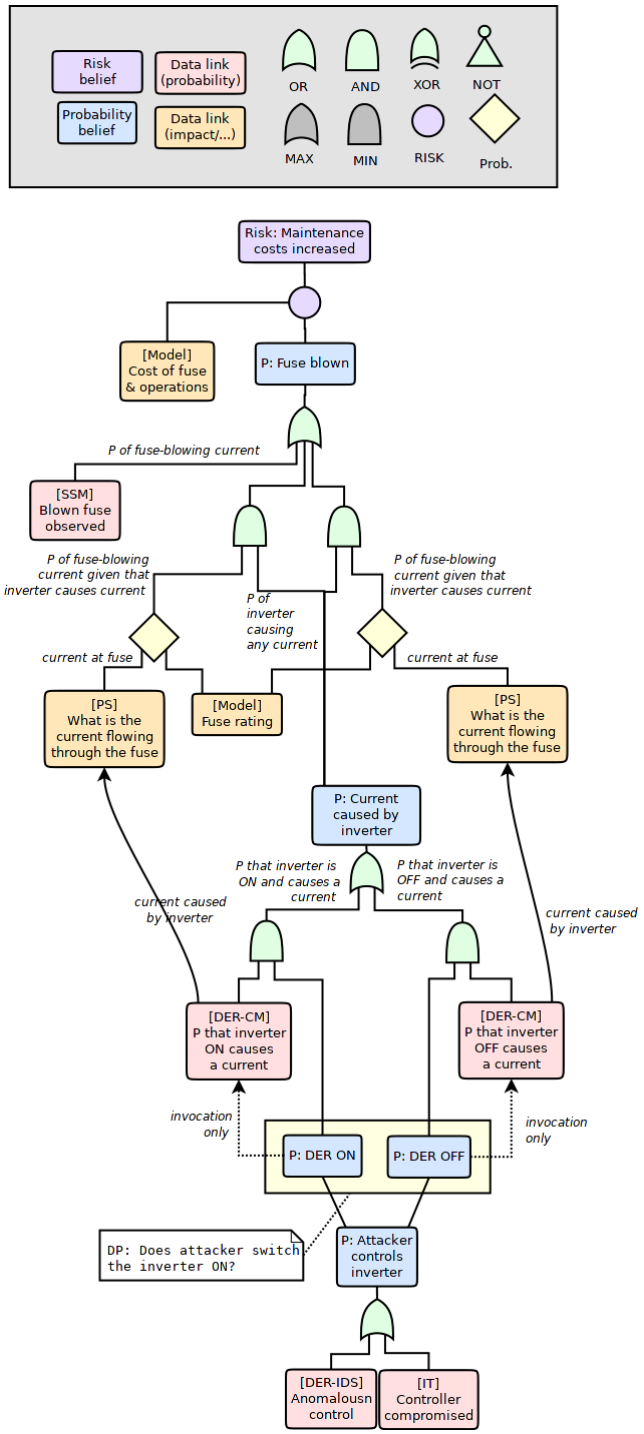

Figure 4: Hypothesis template for the fuse blow hypothesis

and data obtained by executing one of the domain-specific tools (e.g. a power system impact analysis). At each point of the tree traverse where such external input is needed, a data request is forwarded to the tool orchestrator module which coordinates the execution of one or more domain-specific tools, prepares the input data for these tools and collects their output. The execution of each individual tool is encapsulated into an executor module to isolate the rest of the framework from tool-specific execution modes such as shell invocations, platform dependencies etc. Once the tree has been traversed, the value of the root node is passed on to the ranking module, which applies the ranking criteria provided by the hypothesis generator. The ranking module will only terminate when the last hypothesis has been generated and resolved. Its output is a list of (the highest ranking) hypothesis ordered by their risk value.

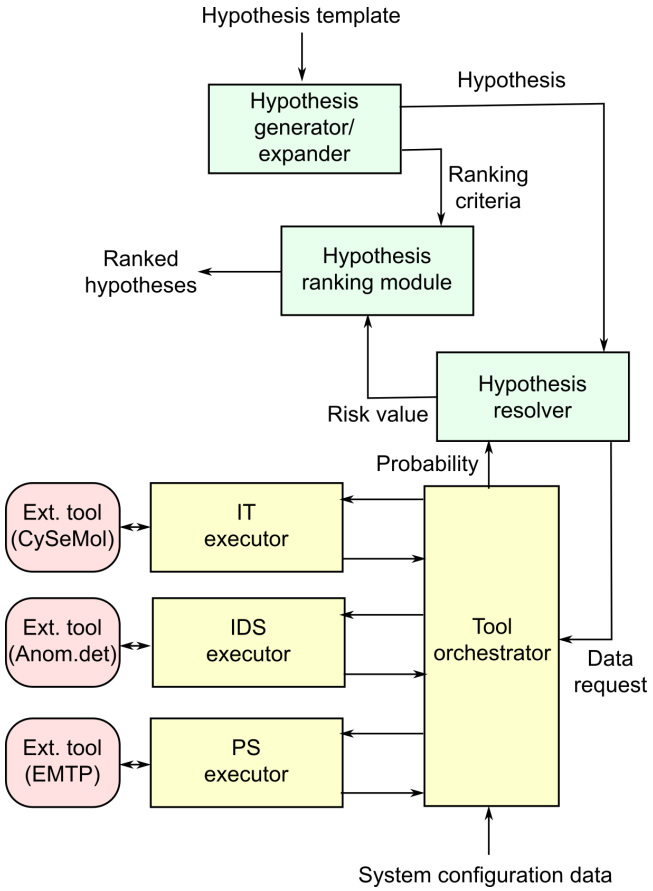

Figure 5: Framework architecture and data flow

6 CONCLUSION

In the above paper we have discussed and presented an approach for the integration of cybersecurity tools from multiple domains into an overall risk assessment tool which takes the complex interactions between domains in smart grid systems into account. We have also presented a very simple example case to serve as a proof of concept of the chosen approach, as well as a generic software framework for the processing of hypothesis templates. Both the solution approach and the software implementing it are currently at an early stage of development. The next logical development step is the application of the method to a more complex and more

realistic scenario. A case has been developed in the SALVAGE project which investigates possible attacks on the configuration of protection devices in a medium voltage grid. The creation of an appropriate hypothesis template for this system is planned. This template will include several branching points, allowing the investigation of practical methods for mitigating the "explosion of hypotheses". The applicability of the presented approach will to a significant degree depend on solutions for this issue. Further down the road, the partial automation of hypothesis template design presents an interesting challenge. Currently, hypothesis templates have to be hand-crafted by a group of domain experts covering all domains. Identifying some low-hanging fruits to reduce the effort would strengthen the viability of the approach.

ACKNOWLEDGMENTS

This work has received funding from ERA-Net SmartGrids through the Danish ForskEL programme (Grant No. 12254) and the Swedish Centre for Smart Grids and Energy Storage (SweGRIDS).

REFERENCES

[1] CEN-CENELEC-ETSI Smart Grid Coordination Group. 2012. Smart Grid Information Security. (November 2012).

[2] V. Chandola, A. Banerjee, and V. Kumar. 2009. Anomaly detection: A survey. *ACM computing surveys (CSUR)* 41, 3 (2009), 15.

[3] Thomas M Chen, Juan Carlos Sanchez-Aarnoutse, and John Buford. 2011. Petri net modeling of cyber-physical attacks on smart grid. *IEEE Transactions on Smart Grid* 2, 4 (2011), 741–749.

[4] Hannes Holm, Khurram Shahzad, Markus Buschle, and Mathias Ekstedt. 2015. P2CySeMoL: Predictive, Probabilistic Cyber Security Modeling Language. *IEEE Transactions on Dependable and Secure Computing* 12 (Nov.-Dec 2015), 626–639.

[5] Y. Isozaki, S. Yoshizawa, Y. Fujimoto, H. Ishii, I. Ono, T. Onoda, and Y Hayashi. 2014. On detection of cyber attacks against voltage control in distribution power grids. In *Smart Grid Communications (SmartGridComm), 2014. Proceedings. 2014 IEEE Int. Conf. on.* IEEE, 842–847.

[6] Barbara Kordy, Ludovic Piètre-Cambacédès, and Patrick Schweitzer. 2014. DAG-based attack and defense modeling: Don't miss the forest for the attack trees. *Computer science review* 13 (2014), 1–38.

[7] Sudha Krishnamurthy, Soumik Sarkar, and Ashutosh Tewari. 2014. Scalable anomaly detection and isolation in cyber-physical systems using bayesian networks. In *ASME 2014 Dynamic Systems and Control Conference.* American Society of Mechanical Engineers, V002T26A006–V002T26A006.

[8] E. LeMay, M. D. Ford, K. Keefe, W. H. Sanders, and C. Muehrcke. 2011. Model-based Security Metrics Using ADversary VIew Security Evaluation (ADVISE). In *Proc. Eighth Int. Conf. Quantitative Evaluation of SysTems.* 191–200.

[9] K. Mets, J. A. Ojea, and C. Develder. 2014. Combining Power and Communication Network Simulation for Cost-Effective Smart Grid Analysis. *IEEE Communications Surveys Tutorials* 16, 3 (Third 2014), 1771–1796. DOI:http://dx.doi.org/10.1109/SURV.2014.021414.00116

[10] Yilin Mo, Tiffany Hyun-Jin Kim, Kenneth Brancik, Dona Dickinson, Heejo Lee, Adrian Perrig, and Bruno Sinopoli. 2012. Cyber–physical security of a smart grid infrastructure. *Proc. IEEE* 100, 1 (2012), 195–209.

[11] National Technical Authority for Information Assurance. 2009. HMG IA Standard No. 1 Technical Risk assessment. Cheltenham, United Kingdom: National Technical Authority for Information Assurance. (2009).

[12] Xinming Ou, Sudhakar Govindavajhala, and Andrew W Appel. 2005. MulVAL: A Logic-based Network Security Analyzer. In *USENIX security.*

[13] Fabio Pasqualetti, Florian Dörfler, and Francesco Bullo. 2013. Attack detection and identification in cyber-physical systems. *IEEE Trans. Automat. Control* 58, 11 (2013), 2715–2729.

[14] M. Sanz-Bobi, A.M. San Roque, A. de Marcos, and M. Bada. 2012. Intelligent system for a remote diagnosis of a photovoltaic solar power plant. *Journal of Physics: Conference Series* 364, 1 (2012).

[15] Oleg Sheyner, Joshua Haines, Somesh Jha, Richard Lippmann, and Jeannette M Wing. 2002. Automated generation and analysis of attack graphs. In *Security and privacy, 2002. Proceedings. 2002 IEEE Symposium on.* IEEE, 273–284.

[16] Teodor Sommestad, Mathias Ekstedt, and Hannes Holm. 2013. The Cyber Security Modeling Language: A Tool for Assessing the Vulnerability of Enterprise System Architectures. *Systems Journal, IEEE* 7, 3 (Sept 2013), 363–373.

[17] Siddharth Sridhar, Adam Hahn, and Manimaran Govindarasu. 2012. Cyber–physical system security for the electric power grid. *Proc. IEEE* 100, 1 (2012), 210–224.

[18] The Open Group. 2013. Open Group Standard: Risk Analysis (O-RA. Berkshire, United Kingdom: The Open Group. (2013). https://www2.opengroup.org/ogsys/catalog/C13G

[19] An Hoa Vu, Nils Ole Tippenhauer, Binbin Chen, David M. Nicol, and Zbigniew Kalbarczyk. 2014. CyberSAGE: A Tool for Automatic Security Assessment of Cyber-Physical Systems. In *Quantitative Evaluation of Systems.* Springer.

[20] D. Yang, A. Usynin, and J.W. Hines. 2006. Anomaly-based intrusion detection for SCADA systems. In *Nuclear Plant Instrumentation Controls and Human Machine Interface Technology, 2006. Proceedings. 5. Intl. Topical Meeting on.* American Nuclear Society, 797–803.

[21] Tim Yardley, Robin Berthier, David Nicol, and William H Sanders. 2013. Smart grid protocol testing through cyber-physical testbeds. In *Innovative Smart Grid Technologies (ISGT), 2013 IEEE PES.* IEEE, 1–6.

[22] A. Zaher, S. McArthur, D. Infield, and Y. Patel. 2009. Online wind turbine fault detection through automated SCADA data analysis. *Wind Energy* 12, 6 (2009), 574.

[23] Kim Zetter. 2016. Inside the cunning, unprecented hack of Ukraine's power grid. (2016). https://www.wired.com/2016/03/inside-cunning-unprecedented-hack-ukraines-power-grid/